

A URA deixou de ser uma árvore de menus

Como o NIVA e o SipPulse AI resolvem chamada de verdade: blocos visuais no lugar do digite 1, agentes de IA com acesso aos seus sistemas e modelos de fala treinados para o áudio que a telefonia realmente entrega.

Equipe SipPulse NIVA, SipPulse AI sippulse.com

O ESSENCIAL

A URA tradicional roteia a chamada. O NIVA resolve a demanda, e mais de 60% do primeiro nível encerra sem agente humano.

Os fluxos são planos visuais montados em 21 blocos. Quem opera o atendimento edita e publica, sem release de TI.

O SipPulse AI é a camada de inteligência: agentes com ferramentas, modelos próprios de telefonia e análise de qualidade em toda a base.

O NIVA licencia por canal simultâneo, não por minuto de contato. O consumo de IA tem teto por sessão e cota semanal.

Toda operação de atendimento conhece a cena. O cliente liga, navega pelo menu, digita o CPF, escolhe a opção, espera. Quando o agente humano finalmente atende, a primeira frase é sempre a mesma pergunta: me confirma o seu CPF, por favor.

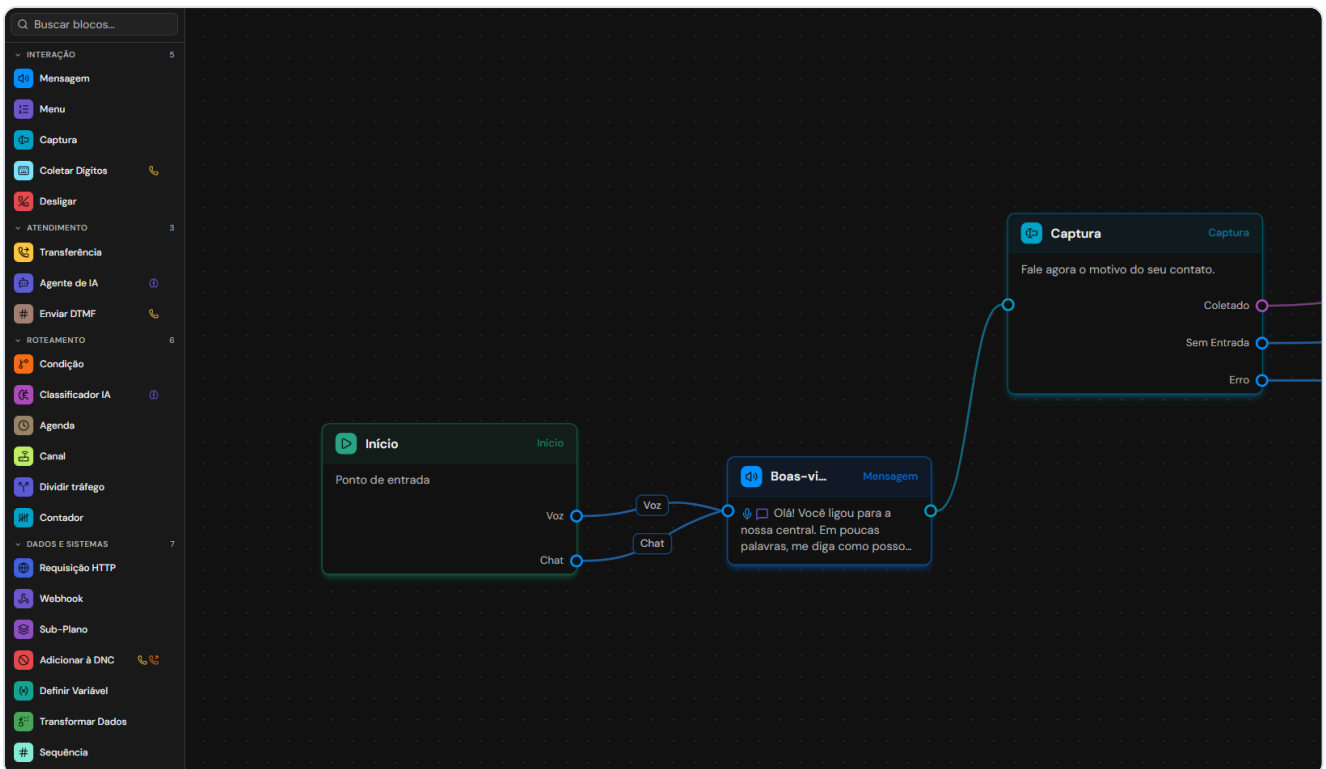
Esse é o limite estrutural da URA tradicional. Ela foi desenhada para **encaminhar** a chamada, não para resolvê-la. A árvore de menus é boa em colocar o cliente numa fila e ruim em qualquer outra coisa: não entende o motivo real da ligação, não consulta o seu sistema, não guarda o que já foi dito. E mudar uma frase do script depende de ticket para TI e janela de release.

O NIVA nasceu para atacar os três problemas ao mesmo tempo. É a URA da SipPulse reconstruída como blocos visuais, com IA de ponta a ponta e integração com os sistemas que a sua operação já usa.

Um plano visual no lugar do fluxograma na parede

No NIVA cada etapa da chamada é um bloco. Atender, escutar, consultar uma API, decidir o caminho, falar, transferir. O plano de atendimento é a ligação entre esses blocos, montada arrastando na tela.

O detalhe que muda a operação está nas saídas. O bloco Captura, por exemplo, não entrega um caminho único: ele expõe três, para coletado, sem entrada e erro. Nada de comportamento implícito escondido em código, o que significa que a pessoa que desenhou o atendimento consegue ler o plano inteiro e enxergar onde o cliente se perde.



O editor de planos do NIVA. A paleta à esquerda traz os blocos por família, e o bloco Início já separa as saídas de Voz e de Chat, então o mesmo plano atende os dois canais.

São 21 blocos organizados em quatro famílias. Interação cobre mensagem, menu, captura e coleta de dígitos. Atendimento traz transferência, envio de DTMF e o bloco Agente de IA. Roteamento tem condição, agenda, canal, divisão de tráfego, contador e o Classificador IA. Dados e Sistemas fecha com requisição HTTP, webhook, sub-plano, definição e transformação de variáveis, sequência e inclusão em lista de bloqueio.

Vale isolar dois deles, porque são os que fazem o produto sair da categoria URA.

Classificador IA

Decide o caminho pela intenção que o cliente declara falando, sem árvore de menu. O roteamento passa a acompanhar o que a pessoa quer, não o que ela conseguiu digitar.

Agente de IA

Entrega a conversa a um agente do SipPulse AI e devolve o resultado ao plano. É a ponte entre a orquestração da chamada e a camada de inteligência.

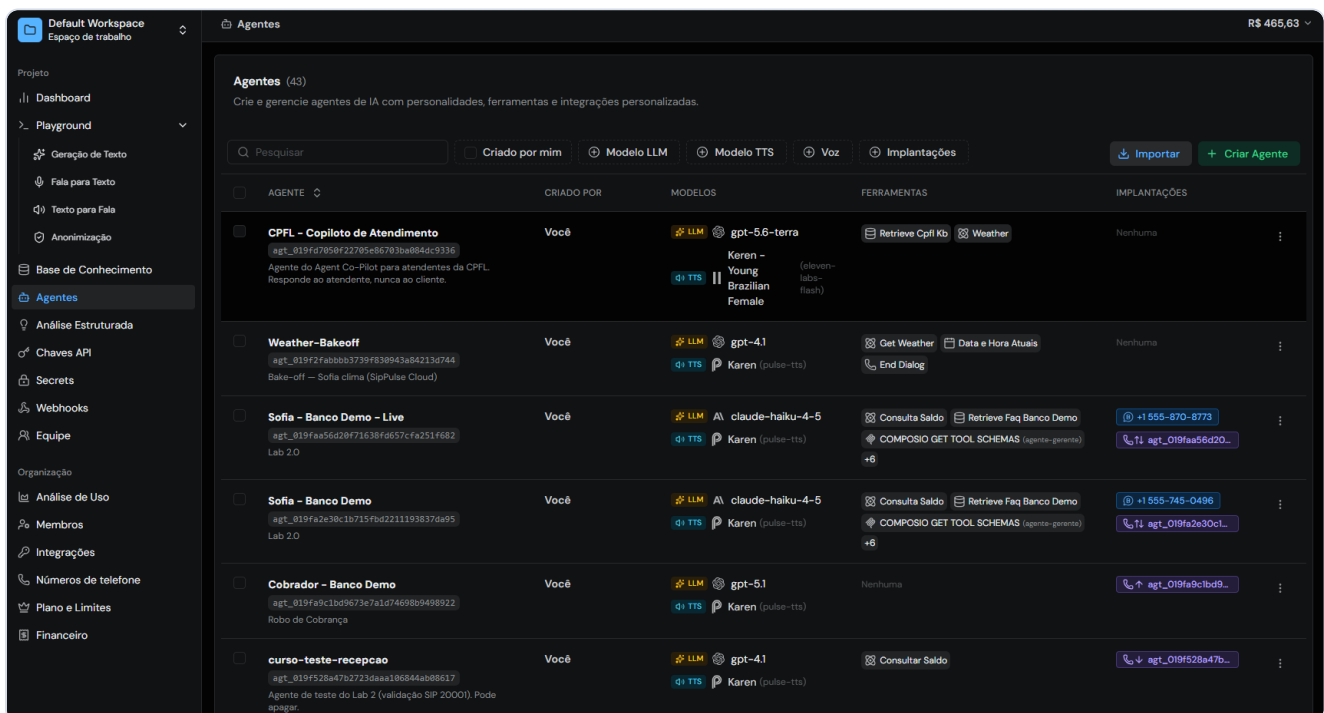
Duas camadas, um fornecedor

Aqui costuma aparecer a confusão, então vale ser explícito. NIVA e SipPulse AI são dois produtos com papéis distintos.

O **NIVA** fica no caminho da voz. Ele cuida da sinalização SIP, do plano visual, das variáveis e condições, do roteamento, do discador, do handover e das integrações. É a camada de orquestração, o produto que a operação abre todo dia.

O **SipPulse AI** é a camada de inteligência. É por ela que o plano acessa os agentes de IA, o reconhecimento de fala, a síntese de voz e os modelos de linguagem. Quando o bloco Agente de IA dispara, é um agente do SipPulse AI que assume a conversa.

Na prática isso significa que você pode começar por qualquer uma das duas. Quem já opera a URA ganha os agentes sem trocar de plataforma. Quem começa pelos agentes sobe para a orquestração completa quando o fluxo cresce.



Default Workspace Espaço de trabalho		Agentes		R\$ 465,63	
Projeto					
Dashboard					
Playground					
Geração de Texto					
Fala para Texto					
Texto para Fala					
Anonimização					
Base de Conhecimento					
Agentes					
Análise Estruturada					
Chaves API					
Secrets					
Webhooks					
Equipe					
Organização					
Análise de Uso					
Membros					
Integrações					
Números de telefone					
Plano e Limites					
Financeiro					
Agentes (43)					
Crie e gerencie agentes de IA com personalidades, ferramentas e integrações personalizadas.					
Pesquisar					
Criado por mim					
Modelo LLM					
Modelo TTS					
Voz					
Implantações					
Importar					
Criar Agente					
AGENTE	CRIADO POR	MODELOS	FERRAMENTAS	IMPLANTAÇÕES	
☐ CPFL - Copiloto de Atendimento agt_019f27050f22705e86703ba084dc9336 Agente do Agent Co-Pilot para atendentes da CPFL. Responde ao atendente, nunca ao cliente.	Você	LLM gpt-5.6-terra Keren - Young Brazilian Female (labov-labov-fresh)	Retrieve Cpf Kb Weather	Nenhuma	
☐ Weather-Bakeoff agt_019f2f60b0a3729f830943a84213d744 Bake-off - Sofia clima (SipPulse Cloud)	Você	LLM gpt-4.1 TTS Karen (pulse-tts)	Get Weather Data e Hora Atual End Dialog	Nenhuma	
☐ Sofia - Banco Demo - Live agt_019fa55d20f71638fd657cfa251f682 Lab 2.0	Você	LLM AI claude-haiku-4-5 TTS Karen (pulse-tts)	Consulta Saldo Retrieve Faq Banco Demo COMPOSIO GET TOOL SCHEMAS (agente-gerencia) +6	+1 555-870-8773 agt_019fa55d20f71638fd657cfa251f682	
☐ Sofia - Banco Demo agt_019fa2e30c1b715fbd2211193837d495 Lab 2.0	Você	LLM AI claude-haiku-4-5 TTS Karen (pulse-tts)	Consulta Saldo Retrieve Faq Banco Demo COMPOSIO GET TOOL SCHEMAS (agente-gerencia) +6	+1 555-745-0496 agt_019fa2e30c1b715fbd2211193837d495	
☐ Cobrador - Banco Demo agt_019fa9c1bd9b73e7a1d7469809498922 Robo de Cobrança	Você	LLM gpt-5.1 TTS Karen (pulse-tts)	Nenhuma	agt_019fa9c1bd9b73e7a1d7469809498922	
☐ curso-teste-recepcao agt_019f52847b2723daa18684a0b88617 Agente de teste do Lab 2 (validação SIP 2000). Pode apagar.	Você	LLM gpt-4.1 TTS Karen (pulse-tts)	Consultar Saldo	agt_019f52847b2723daa18684a0b88617	

O console de agentes do SipPulse AI. Cada agente escolhe o próprio modelo de linguagem e a própria voz, e a lista mostra as ferramentas conectadas e as implantações ativas por agente.

Um agente que só conversa é demonstração

O que separa um piloto bonito de um agente em produção é a lista de ferramentas. Sem elas, o agente responde bem e resolve nada.

No SipPulse AI o agente pode chamar o seu endpoint durante a conversa, com as credenciais guardadas em cofre de secrets fora do prompt. Pode consultar uma base de conhecimento montada com os seus documentos, em vez de responder pelo que o modelo acha que sabe. Traz as funções nativas de telefonia já prontas, como transferir, encerrar e enviar DTMF. E conecta catálogos inteiros de ferramentas por servidores MCP, sem escrever um wrapper para cada integração.

Do lado da chamada, o agente se comporta como um ponto SIP de verdade: atende a fila e também origina ligação, transfere para o ramal certo, recebe dígitos do cliente e envia dígitos para URAs de terceiros, o que costuma ser o que destrava a integração com sistema legado. Ao encerrar, entrega por webhook o que foi coletado e o resultado da análise para o seu CRM ou BI.

Modelo genérico erra no áudio de telefonia

Reconhecimento de fala em condição de laboratório é problema resolvido. Reconhecimento de fala em ligação telefônica não é. O áudio chega em 8kHz, com ruído de linha, sotaque, conversa cruzada e cliente falando de dentro do carro. Modelo treinado em podcast tropeça exatamente aí.

Por isso a SipPulse mantém modelos próprios em vez de repassar engine de terceiro.

RECONHECIMENTO DE FALA

pulse-telephony

Modelo especializado em reconhecimento de voz para telefonia. Também roda em streaming, para co-piloto de agente e inteligência conversacional, com execução on-premises.

SÍNTESE DE VOZ

pulse-tts

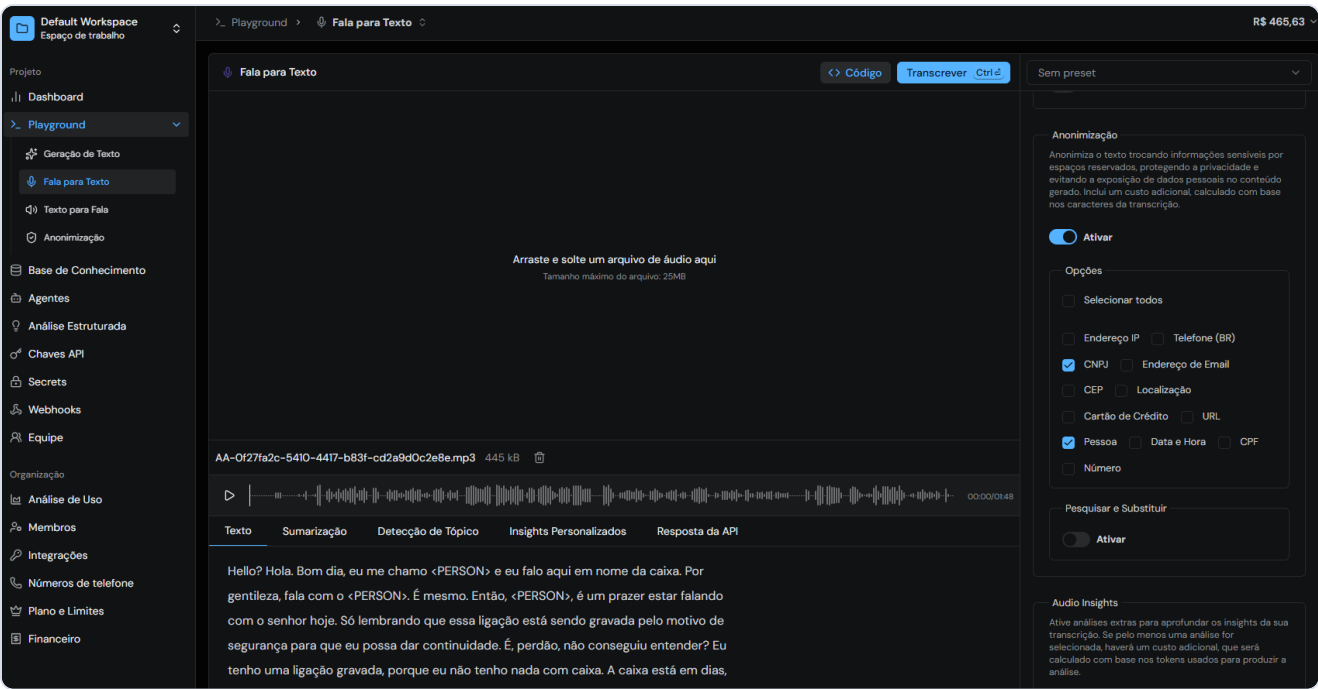
TTS de baixa latência para agentes de voz, com possibilidade de execução local quando a latência ou o dado do cliente exigem.

Quando a expressividade da voz é o critério em vez da latência, a plataforma também dá acesso a modelos de fala em estado da arte como ElevenLabs e FishAudio pelo mesmo endpoint. Você escolhe a voz por caso de uso, não por contrato.

Voz é dado biométrico, e isso muda o projeto

Em banco, utility e operadora a conversa sobre IA de voz chega rápido na LGPD, porque a gravação de voz é dado biométrico. A resposta da plataforma é remover a informação sensível na própria transcrição, antes que ela circule.

CPF, CNPJ, telefone, cartão, e-mail, CEP, endereço IP, localização e nome de pessoa saem da transcrição por seleção. Para quem precisa de garantia mais forte, o pulse-telephony em streaming e o pulse-tts rodam on-premises, e aí o áudio do cliente não sai da sua infraestrutura.

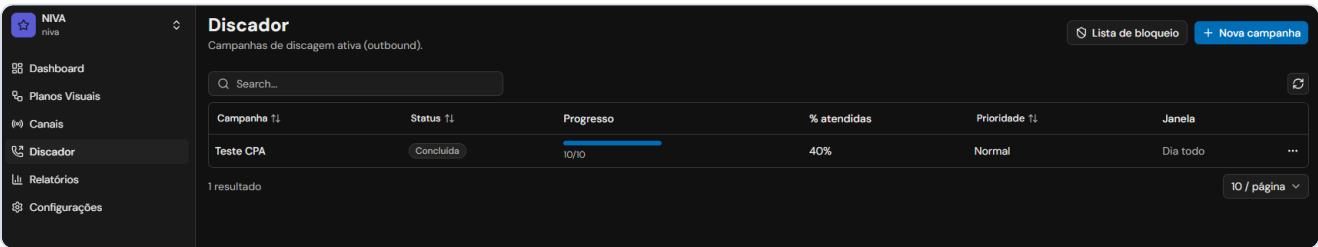


O playground de fala para texto, com o painel de anonimização à direita. Além de testar modelo e parâmetros no navegador, ele gera o código equivalente em cURL, Python e JavaScript.

O NIVA também disca

Automação de atendimento costuma ser discutida só no inbound, mas a operação de cobrança, pesquisa e confirmação de agendamento vive do outbound. O discador é nativo do NIVA, com campanhas que têm prioridade, janela permitida de discagem e progresso acompanhado.

Duas peças fazem diferença no custo dessa operação. O classificador CPA distingue pessoa de caixa postal com 99,9% de acurácia, o que evita que o agente humano receba secretária eletrônica. E a lista de bloqueio é alimentada pelo próprio plano, pelo bloco Adicionar à DNC, durante a conversa em que o cliente pede para não ser mais contatado.



Campanhas de discagem no NIVA, com prioridade, janela permitida, progresso e percentual atendido por campanha.

Qualidade em toda a base, não na amostra

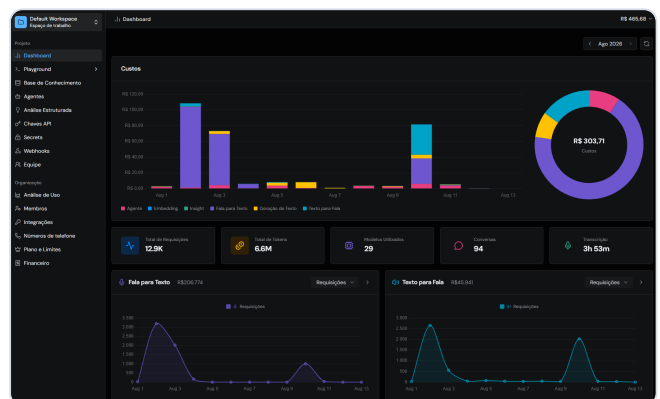
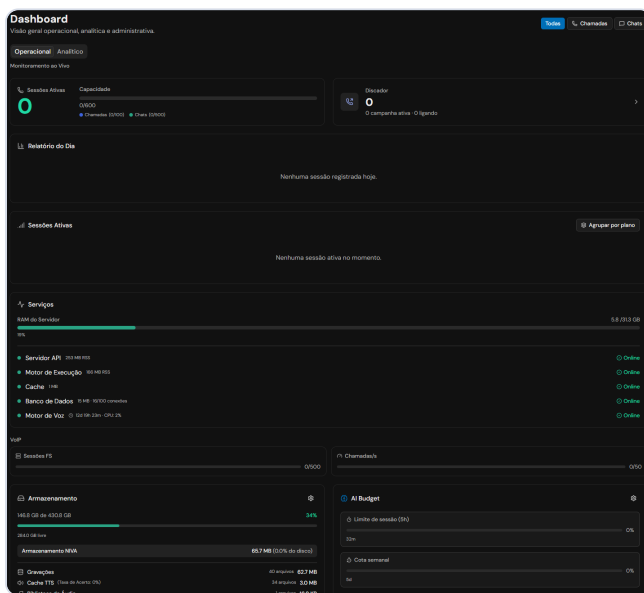
A monitoria tradicional ouve o que cabe no dia do supervisor. Na prática, uma fração pequena das ligações define a nota de todo mundo.

A análise estruturada do SipPulse AI inverte isso. Você declara o schema no construtor visual de campos, transformando cada critério em um campo com tipo de retorno: saudação, sondagem, oferta correta, script de compliance, encerramento. A plataforma aplica o mesmo critério em todas as ligações e todos os chats, e devolve os campos preenchidos por conversa, junto de sentimento, tópico e intenção.

O resultado sai por webhook para o seu painel assim que fica pronto, e o mesmo schema pode ser exportado entre times e projetos, o que mantém o critério igual em voz e em chat, em agente de IA e em atendimento humano.

Onde o custo aparece

A objeção mais comum a projeto de IA de voz não é técnica, é de previsibilidade: quanto isso vai custar no mês em que o volume dobrar. As duas camadas respondem de formas diferentes, e as duas expõem o número.



À esquerda, o dashboard operacional do NIVA: capacidade de sessões simultâneas, saúde dos serviços e cache de TTS. À direita, o console do SipPulse AI, com custo separado por serviço e cobrança em real.

NIVA: por canal simultâneo

A licença acompanha o pico de sessões simultâneas, contadas separadamente para chamadas e chats. Não há cobrança por minuto de contato nem por volume mensal, então uma conversa mais longa não muda a conta. Disponível on-premise ou em cloud.

SipPulse AI: créditos em real, com teto

Consumo em créditos com impostos inclusos, e cada componente com sua unidade: minuto de áudio no reconhecimento de fala, caractere na síntese de voz, token na geração de texto. O AI Budget define limite por sessão e cota semanal, e o dashboard mostra o gasto por serviço dia a dia.

Por onde começar

Projeto de automação de atendimento não começa escolhendo modelo. Começa olhando o relatório de chamadas e separando o que é demanda repetitiva com resposta objetiva, porque é ali que a automação paga rápido.

A partir desse recorte, o caminho costuma ser o mesmo: mapear o fluxo e as integrações na primeira semana, montar os planos nos blocos e conectar as APIs nas semanas seguintes, e colocar um piloto em tráfego real sobre uma fatia controlada da fila antes de ampliar. Os relatórios e as métricas de automação passam a dizer qual bloco ajustar na semana seguinte.

Quer ver o NIVA rodando com os seus dados?

Trinta minutos com o seu relatório de chamadas em mão. Saímos com o percentual automatizável, o desenho dos planos e o número de canais simultâneos para dimensionar.

sippulse.com docs.sippulse.ai info@sippulse.com +55 48 3332-8540